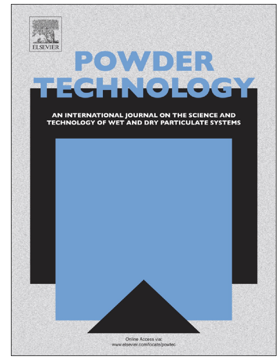
Developing ANN-Kriging hybrid model based on process parameters for prediction of mean residence time distribution in twin-screw wet granulation

Article in Powder Technology · November 2018

DOI: 10.1016/j.powtec.2018.11.060

Accepted Manuscript



Developing ANN-Kriging hybrid model based on process parameters for prediction of mean residence time distribution in twin-screw wet granulation

Hamza Y. Ismail, Mehakpreet Singh, Shaza Darwish, Manuel Kuhs, Saeed Shirazian, Denise M. Croker, Majeda Khraisheh, Ahmad B. Albadarin, Gavin M. Walker

PII: S0032-5910(18)30964-1
DOI: https://doi.org/10.1016/j.powtec.2018.11.060
Reference: PTEC 13895

To appear in: *Powder Technology*

Received date: 14 July 2018
Revised date: 24 October 2018
Accepted date: 12 November 2018

Please cite this article as: Hamza Y. Ismail, Mehakpreet Singh, Shaza Darwish, Manuel Kuhs, Saeed Shirazian, Denise M. Croker, Majeda Khraisheh, Ahmad B. Albadarin, Gavin M. Walker , Developing ANN-Kriging hybrid model based on process parameters for prediction of mean residence time distribution in twin-screw wet granulation. Ptec (2018), https://doi.org/10.1016/j.powtec.2018.11.060

Developing ANN-Kriging hybrid model based on process parameters for prediction of mean residence time distribution in twin-screw wet granulation

Hamza Y. Ismail[1,2], Mehakpreet Singh[2], Shaza Darwish[2], Manuel Kuhs[2], Saeed Shirazian[2,*], Denise M. Croker[2], Majeda Khraisheh[3], Ahmad B. Albadarin[1,2], Gavin M. Walker[1,2]

[1] Pharmaceutical Manufacturing Technology Centre, Bernal Institute, University of Limerick, Limerick, Ireland

[2] Department of Chemical Sciences, Bernal Institute, University of Limerick, Limerick, Ireland

[3] Department of Chemical Engineering, College of Engineering, Qatar University, Qatar

[*] Corresponding author: saeed.shirazian@ul.ie

Tel.: +353 (0) 61 23 7718

Abstract

Artificial neural network (ANN) modelling is applied to predict the mean residence time of pharmaceutical formulation in a twin-screw granulator. Process parameters including feed flow rate, screw speed, and liquid to solid ratio are correlated with the obtained values of mean residence time to build a predictive tool. In order to improve the ANN predictive capability, a kriging interpolation approach is utilised and both ANN models (before and after kriging) are compared. Experimental data is obtained for wet granulation of microcrystalline cellulose using a bench-scale 12 mm twin-screw granulator. In addition, the effect of screw configurations on mean residence time is investigated by the developed ANN. The ANN model is made of two hidden layers with 2 linear nodes in each layer, and the linear system of equations is derived for the improved ANN model. The results revealed that the developed model was capable of predicting the mean residence time in the granulator more accurately after applying kriging interpolation, with an R^2 value of about 0.92 for both training and validation. ANN model after kriging shows a dramatic improvement of R^2 by 4% and 22% in training and validating phases, respectively. Also, the RMSE was improved by 40% and 61.5% in training and validating phases, respectively. Furthermore, this improvement was reflected in the contour profiles of the ANN models before and after kriging interpolation, where the model that uses the interpolated data points shows a smoother contour profiles and wider prediction areas. Screw configuration has the most significant effect on the residence time of granules inside the granulator where adding more kneading zones results in a substantial increase in the mean residence time compared to other process parameters.

Keywords: Artificial neural network; Kriging; Twin-screw granulator; Continuous pharmaceutical manufacturing; Residence time; Model predictive control

Abbreviations

ANN	artificial neural network
Ce	conveying elements
Cov	covariance
Cut	cutting elements
DEM	discrete element method
DoE	design of experiment
GP	genetic programming
HDMR	Higher dimensional model representation
Kz	kneading zone
MCC	microcrystalline cellulose
MRT	Mean residence time
PBM	population balance model
QbD	Quality-by-Design
QbT	Quality-by-Testing
RMSE	root-mean-squared error
RSM	Response surface methodology
RTD	residence time distribution
SSE	sum of squared error
TSG	Twin-screw granulation

1. Introduction

The pharmaceutical industry has traditionally relied on batch processing for manufacture despite its many disadvantages. Batch processing is wasteful, time-consuming, offers poor control over product quality, and is high in cost[1]. Efficiency wise, a batch that does not meet the quality requirements goes entirely to waste due to the use of Quality-by-Testing (QbT) method in product analysis [2]. Recently, the pharmaceutical industry has become increasingly interested in process modelling and understanding in order to move from batch to continuous processing. Various models including mechanistic and data-driven models have been developed to describe and understand the pharmaceutical processing [3-6]. The key processing step in the pharmaceutical manufacture of solid-dosage formulations is granulation in which granules are produced from fine powder including formulations [7, 8].

In order to correlate the granule properties to process parameters, process modelling is required. There are various process models for the design and optimisation of granulation such as mechanistic and data-driven models. The major mechanistic models for the design of granulation processes include the population balance model (PBM) and the discrete element method (DEM). PBM is based on the conservation of the number density of particles and tracks the particle properties during granulation, while DEM is based on Newton's second law and tracks individual particles in the spatial coordinate [9, 10]. There are various applications of PBM and DEM for modelling of wet granulation in the literature [11-15]. However, there are some disadvantages associated with these mechanistic models, given the fact that they are computationally expensive to use for commercial applications. Therefore, more robust models are required for development of a Quality-by-Design (QbD) approach in pharmaceutical manufacturing. The model should be fast and accurate enough to capture the granule properties, such as size, composition, porosity etc. [3].

Data-driven models have been shown to be of great potential for simulation of pharmaceutical processes. Recently, our research group developed the idea of data-driven models for simulation of dry granulation by roller compactor [4]. They used artificial neural network (ANN) modelling for prediction of ribbon density obtained in a roller compactor. Their results indicated that the ANN is a great tool for predicting roller compactor, and the product properties can be correlated very well to the process parameters as well as input material properties. Kazemi et al. [16] simulated oscillating milling used for pharmaceutical formulations with the aid of different intelligence models. They utilised generic programming (GP) and ANN in order to predict the granule size during the milling step. Their results revealed that ANN and GP are powerful tools for simulation of pharmaceutical processes. Other researchers have investigated the application of ANN for simulation and optimisation of wet granulation, and concluded that these data-driven models are robust and reliable in descripting pharmaceutical processes [5, 17].

It has been recognised that modelling using an ANN method requires large amounts of measured data for training and testing which causes additional costs. In order to reduce the costs of measurement for developing an ANN model, one can use kriging to generate more data points. Kriging is an interpolation method in which the values are modelled by Gaussian process, and it has proved to be a powerful empirical tool in modelling complex processes. Kriging has been successfully used recently in pharmaceutical applications such as predicting the ribbon density and roll gap in a roller compactor in terms of hydraulic pressure and roller speed. Boukouvala et al. [18] showed that kriging overtook the ANN model for simulating dynamic processes in pharmaceutical manufacturing. In another work, Boukouvala et al. [19] used kriging and other empirical tools in predicting the residence time distribution of powders in a continuous mixer with respect to impeller rotational rate and powder flow rate, where kriging showed a better performance than the other empirical and interpolation tools including ANN, response surface methodology (RSM), and higher dimensional model representation (HDMR).

Residence time distribution (RTD) is of great importance for design and scale-up of twin-screw wet granulation because the granule properties depend on the duration that particles reside inside the granulator. Various phenomena including nucleation, aggregation, and breakage occur during twin-screw wet granulation and all mechanisms are time dependent [20]. Given the fact that short residence time is the main advantage of twin-screw granulation, solid-liquid mixing should be achieved as quickly as possible. This means an appropriate arrangement of screw configuration as well as optimisation of process parameters are required in order to produce granules with the desired properties in a short time [21]. Although there has been research on residence time distribution in twin-screw granulation, the main focus has been on fitting a correlation to predict RTD [20]. However, determination of mean residence time (MRT) would be of great interest from the practical point of view, as it can be used for predictive modelling of particle size distribution. Development of a robust predictive model for the MRT in granulation process would help optimise the twin-screw wet granulation for pharmaceutical manufacture. To the best of our knowledge, there is no report on hybrid ANN-kriging analysis of residence time in twin-screw wet granulation of pharmaceutical formulations.

The current work aims to develop an improved ANN (hybrid ANN-kriging) model for simulation of the granulation process. Twin-screw wet granulation is studied, and the experimental data are obtained for granulation of pure microcrystalline cellulose. An ANN model is developed using original experimental and interpolated data (after kriging) to correlate the mean residence time of particles with process parameters and screw configurations.

2. Experimental setup

A twin-screw granulator (*ThreeTec*, Switzerland) with diameter of 1.2 cm and length-to-diameter ratio 40:1 was used for granulation of pure microcrystalline cellulose (MCC, Avicel PH 101). Distilled water was added as the liquid binder using a peristaltic pump (Watson Marlow 520SN). MCC flow rate was adjusted by a gravimetric feeder (Three-Tec, Switzerland). The Design of Experiment (DoE) used for granulation experiments is listed in Table 1. A custom-designed method was used to design the experiments, considering feed flow rate, screw speed, liquid to solid ratio (*L/S*), and screw configurations as the factors and mean residence time as the response. The parameter levels in DoE were chosen based on the process limitation and preliminary experiments. Liquid to solid ratio varied between 0.54 and 1.22 because the granules were formed in this range. Moreover, the screw speed varied between 50 to 200 rpm. Fig. 1 illustrates the experimental setup used in this work. Four screw configurations were chosen: conveying zone only (Ce), one kneading zone (1Kz) consisting of six kneading elements at 60° angle, two kneading zones (2Kz) consisting of a kneading zone of six elements separated by a conveying zone from another kneading zone with five kneading elements, and lastly two kneading zones with cutting elements (2Kz+Cut) configuration with addition of cutting elements with adjacent reverse at the end (see supplementary file).

Fig. 1: Photo of twin-screw extruder used for the experiments.

Digital imaging method similar to the one used by Kumar et al. [22] and Mu et al. [23] was used for residence time distribution (RTD) measurements. The methylene blue dye was added as a pulse at the inlet of the extruder. A digital video camera mounted at the outlet of the granulator was used to record the granules exiting the granulator against the white background of the conveyor belt. The intensity of the colour was then plotted as a function of time. The results were fitted using the Zusatz function described in Eq. (1) [24]:

$$E(t) = at^{-c-1}b^{c+1}\exp[(b^c t^{-c} - 1)\left(\frac{-c-1}{c}\right)] \quad (1)$$

where a corresponds to the peak height, b corresponds to the residence time at the peak height, and c is a parameter related to the peak breadth. The exit age distribution was derived from the experimental results and the mean residence time calculated using Eq. (2) and Eq. (3) [24]:

$$E(t) = \frac{C(t)}{\int_0^\infty C(t)dt} \quad (2)$$

where $C(t)$ is the dye concentration at time t, in this case the intensity of the dye.

$$\bar{t} = \frac{\int_0^\infty t.E(t)dt}{\int_0^\infty E(t)dt} \quad (3)$$

3. Modelling approach

ANN modelling is used as an empirical model to predict the mean residence time of granules in twin-screw in terms of the process parameters and screw configuration. Due to limitations in the number of experimental points available for training and validating the ANN model, kriging interpolation is used to interpolate the mean residence time on new data points and use the new interpolated points to build an improved ANN model. Both models are trained and then validated. Kriging interpolation builds a statistical graphical variogram based on the experimental points and then uses the fitted variogram to interpolate the output at the new input points. In this way more points can be used to train and validate a more stable and accurate ANN model that better represents the system.

An ANN model consists of input, hidden and output layers that should be designed in a proper manner [16]. Input and output layers refer to process parameters (factors) and granule properties (responses) respectively. The hidden layer contains the correlation between inputs and the output by either linear or non-linear combination of parameters but linear nodes are usually preferred, if available, because they can be used as state space for model predictive control. In order to develop the optimum ANN for process simulation, a portion of the obtained experimental points are used to train the network, while the other portion of results are used to assess the model, i.e. model validation. Given that there are no straightforward guidelines to find the optimum ANN structure, the number of layers, the number of training points, and the type of activation, a heuristic approach was used for this particular process. *Kfold* method was used for validation of ANN model with K=3. This means that 1/3 of experimental runs are used for model assessment. This validation method searches for different validation possibilities and finds the optimum validation data point with the lowest error. Multi-start option was activated in the ANN model in order to find the global optimum points. *Matlab* was used to conduct the kriging interpolation and *JMP Pro 14* software was used for ANN modelling and analysis of the results [5].

The developed ANN models are used for prediction of mean residence time of microcrystalline cellulose (MCC) granules in a twin-screw granulator, and both models (before and after kriging) were compared by assessing their accuracy in predicting the mean residence time(s). The designed flowchart in Fig. 2 shows the two routes taken in building the ANN model, the first route is using the experimental data and the second route is using the interpolated data after applying kriging interpolation.

Fig. 2: Designed flow chart for building ANN models for predicting the mean residence time in twin-screw granulator of MCC.

4. Results and discussion

4.1. Experimental results

The measured mean residence times of granules inside the granulator are reported in Table 2. As indicated, 24 runs were carried out and the mean residence time for each run was measured according to the method described in Section 2. In order to illustrate the shape of residence time distribution, four examples of RTD for different screw configurations (runs 6, 13, 16 and 24) are given in Fig. 3. The parameters of model RTD (*a*, *b*, *c*) are calculated through comparing with measured values and utilising least squared method. *Matlab* was used for curve fitting and finding the values of unknown parameters. It is shown that the residence times exhibited a wide distribution that could be attributed to the flow regime inside the granulator, as the granulator consists of kneading and conveying zones. Basically, a wide distribution is typical of the TSG [7, 21, 22, 24]. The mean residence time of particles inside the barrel varied from 18.41 to 277.24 seconds depending on the process parameters and screw configuration.

Fig. 3: Residence time distribution of granules samples for different screw configurations.

4.2. Kriging interpolation

Kriging interpolation predicts a response y_k at an interpolated point x_k as a weighted sum of the observed responses $(y_1, y_2 \ldots y_n)$ in which x_k falls in the neighbourhood of their corresponding sampling points $(x_1, x_2, x_3 \ldots x_n)$ [18], where:

$$y_k = f(x_k) = \sum_{i=1}^{n} w_i f(x_i) \tag{4}$$

w_i being the weighted sum (kriging weights) that depends on the Euclidian distance h where

$$h = \|x_i - x_j\| \tag{5}$$

The main objective in the kriging algorithm is to determine the set of kriging weights assigned to each group of n clustered points in the neighbourhood of x_k. The derived variogram model forces the sum of the weights to unity. In determining the interpolated prediction $f(x_k)$ at x_k, function values y_i for sampled points x_i that are in the neighbourhood and nearer to x_k will have more influence on predicting $f(x_k)$. In other words, the higher number of neighbouring points and the nearer these neighbouring points are to x_k, $f(x_k)$ will be predicted with more confidence.

The variogram is derived from the observed data points to statistically quantify the dataset roughness and this complements histograms and other descriptive statistics (exponential, Gaussian, cubic, etc.). The variogram depends on the Euclidian distance h, at each variance difference of the observations in the neighbourhood of x_k, where:

$$\gamma(h) = \frac{1}{2}[var(y_i - y_j)] \tag{6}$$

After fitting the derived variogram to a variogram model the covariance at each h is calculated by:

$$Cov(h) = \sigma_{max}^2 - \gamma(h) \tag{7}$$

σ_{max}^2 is the maximum variance of the variogram function. The covariance is then used to calculate the kriging weight at each point by solving the following system:

$$\begin{bmatrix} Cov(d_{1,1}) & \dots & Cov(d_{1,N}) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ Cov(d_{N,1}) & \dots & Cov(d_{N,N}) & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \times \begin{bmatrix} w_1 \\ \vdots \\ w_N \\ \lambda \end{bmatrix} = \begin{bmatrix} Cov(d_{1,k}) \\ \vdots \\ Cov(d_{N,k}) \\ 1 \end{bmatrix} \quad (8)$$

where $Cov(d_{i,j})$ and $Cov(d_{i,k})$ is the covariance of the distance $d_{i,j}$ and $d_{i,k}$ between sampling points x_i - x_j and x_i - x_k (interpolated point). At each interpolated point x_k the variance is calculated by:

$$\sigma_k^2 = \sigma_{max}^2 - \sum_{i=1}^{N} w_i Cov(d_{i,k}) - \lambda \quad (9)$$

A four-dimensional kriging interpolation was conducted on the experimental data obtained from the twin-screw wet granulation (see Table 1) to predict the mean resident time at new process parameters (liquid to solid ratio, powder feed rate, screw speed and screw configuration).

Kriging was performed in *Matlab* where L/S ratio, screw speed and powder feed rate were taken as input parameters and mean residence time as output parameter at four screw configurations. Interpolation was conducted at 20 new points for each dimension in which 37,045 points were obtained after applying kriging at the four screw configurations in a 3-D numerically meshed system. The interpolated kriging data were used to improve the empirical ANN model prediction compared to using just experimental data. Interpolated data are provided in the supplement excel sheet in Appendix A.

4.3. Artificial neuron network (ANN) modelling

The basic ANN structure consists of three layers: input layer where the input variables x_is are introduced, hidden layers where the data is processed, and an outer layer that represents the output variables y_is. Each hidden layer contains a specific number of nodes; the nodes are made from a range of linear and non-linear equations (*tanh* equation (10), and Gaussian equation (11)):

$$F = \frac{e^{2x} - 1}{e^{2x} + 1} \tag{10}$$

$$F = e^{-x^2} \tag{11}$$

where *F* is the activation function and *x* refers to the linear combination of input variables.

ANN model accuracy is mainly dependent on the number of experimental samples and observation used, the more data points used the more accurate the ANN model becomes. Also, increasing the number of layers and nodes gives the ANN more flexibility in training but makes the validation more difficult and the model more complex; thus it is always preferable to keep the number of layers and nodes used to the minimum and increase the experimental points to obtain a better model. In this work, kriging interpolation is used to generate new data points from the experimental points and thus to improve the ANN model. Simple linear empirical models are always more desirable than nonlinear complex ones especially for control purposes (e.g. model predictive controllers); the disadvantage is that linear equations are less flexible than *tanh* and Gaussian equations. Kriging interpolation is performed on the experimental points and then all the interpolated data are used for training and validating the ANN model using linear nodes. ANN models of experimental data before and after kriging are compared in the following sections. Two hidden layers are considered for ANN development for a better model parsimony, and for each layer, linear activation functions are taken into account. The role of activation functions in ANN is to predict the response values by a combination of input variables (process parameters and screw configuration).

The objective function used in ANN to minimise the error between predicted and measured values may have many local minima which makes the optimisation complex. The absolute minima should be obtained for accurate ANN development. Therefore, a multi-start approach was applied, in which the fitting parameters are estimated by various initial points [5, 16]. The final optimised ANN topology was obtained with two hidden layers, 2 linear nodes in the first and 2 linear nodes in the second layer as represented in Fig. 4. The comparisons between the two ANN models' accuracies (before and after kriging) were carried out in terms of R^2 and RMSE.

Fig. 4: Structure of developed ANN model used for prediction of mean residence time.

4.4. ANN model training and validation before and after kriging

The developed ANN is trained and validated to assess the capability of the model in predicting mean residence time of particles inside the twin screw granulator at various L/S ratio, screw speed, solid feed rate and screw configurations. A number of ANN models were trained and validated for the original experimental data before kriging and for the interpolated data after applying kriging to improve the ANN model prediction accuracy. L/S ratio, screw speed, powder feed rate and screw configuration are the input variables in the first layer, while mean residence time is the output variable in the outer layer as shown in Fig. 4. In both models two thirds of the data points are used for training the ANN model, while the remaining one third are used for validating the trained model.

The ANN is calibrated in the training stage in order to find the weights and bias which are unknown parameters of the model and should be calculated by minimising the objective function. The comparison of models' accuracy in predicting new experimental data for both models (before and after kriging) shows that, after kriging interpolation (see Fig. 5a) the ANN model is more accurate in predicting the mean residence time both in training and in validation compared to the ANN model before kriging (Fig. 5b). Table 3 shows the R^2 improved by about 4% in training and 22% in validation after applying kriging. Also, RMSE was improved by almost 40% for training and 61.5% for validation; similarly, the mean absolute deviation was dramatically improved by 39.3% and 59.6% for training and validation, respectively. This is mainly due to the increased number of points used after applying kriging interpolation, in such a way that the weight of the outliers is light while the weight of the experimental points gathered in the neighbourhoods of each other is higher. The comparisons confirm that kriging interpolation is a powerful tool in improving the ANN model prediction of the mean residence time in the twin-screw granulator. Given that the twin-screw granulation is a complex process on a mechanistic level and many parameters affect the granule properties, the developed ANN is useful understanding the effect of different process parameters on the granule properties.

Fig. 5: Experimental and interpolated data versus predicted mean residence times for ANN models (a) before and (b) after kriging.

4.5. Profiles of the predicted mean residence time in terms of the process parameters and screw configurations

The validated ANN models can be used as a predictive tool to analyse the influence of process parameters and screw configurations on the mean residence time of granules produced by the twin-screw granulator. The profiles for the influence of L/S ratio, screw speed, powder feed rate, and screw configuration on mean residence time for both models (before and after kriging) are shown in Fig. 6. It is shown that the profiles of the process parameters and screw configuration are linearly related to the mean residence time, where increasing L/S ratio increases the mean residence time due to the change in the powder rheological and viscosity making the powder more sticky when moving along the extruder, and hence causing an increase in the mean residence time. On the other hand, it is observed that increasing the screw speed or powder flow rate will decrease the mean residence time. Increasing the screw speed will increase the conveying capacities of the screws resulting in a decrease in the residence time, and higher powder flow rate will cause the powder to move faster inside the granulator resulting in a decrease in the residence time. However, it is seen that the effect of screw speed is more profound, as the main driving force for the movement of granules inside the granulator is the shear force exerted by the screws. Furthermore, the screw configuration profile shows that, for the screw configuration with only convening elements the mean residence time is the shortest, while, for the screw configuration of 2 kneading and 1 cutting element, it is the longest. This is due to the geometry of the conveying elements that transport powder much better than the kneading and the cutting elements leading to a shorter mean residence time. Adding more elements to the screws will result in more hindrance against particle movement, which in turn causes longer residence time.

Fig. 6: Effect of process parameters and screw configurations on mean residence time for (a) before and (b) after kriging.

This trend (relation between the process parameters/screw configuration and the mean residence time) is similar for both ANN models (models before and after kriging) though the steepness of these linear relations decreases for all the profiles for the improved model (after kriging). This decrease is mainly due to the increased number of interpolated points that fall nearer to places where there are more experimental points, giving more weight to the linear function passing through these points, thereby improving the profile prediction accuracy around this group of experimental points.

It is clearly seen from Fig. 6 that kriging interpolation changed the profile of the mean residence time with respect to the process parameters and screw configurations; kriging gives more weight to the group of points in the neighbourhood of each other and less weight to the outliers, thus indirectly giving more weight to the function that relates these nearest experimental points and less weight to the function supported by the outliers. Also, it is clearly seen that screw configuration has an important influence on the mean residence time (see Figs. 7 & 8). The lowest range of mean residence times (40-140 s before kriging, and 60-130 after kriging) is observed for screw with conveying elements only, and the longest is obtained for the screws with two kneading zones and one cutting zone.

The improved ANN model after kriging are represented by a system of empirical linear equations that relate the predicted residence time (s) in terms of L/S ratio, screw speed (rpm), and powder flow rate (g/h) at each screws configuration as the following:

For screw configuration with 1kneading zone:

$$\lambda = \mathbf{127.92 + 55.1\boldsymbol{\alpha} + 0.1179\boldsymbol{\beta} + 0.0822\boldsymbol{\mu}} \tag{12}$$

For screw configuration with 2 kneading zones:

$$\lambda = \mathbf{162.2 + 55.1\boldsymbol{\alpha} + 0.1179\boldsymbol{\beta} + 0.0822\boldsymbol{\mu}} \tag{13}$$

For screw configuration with 2 kneading and 1 cutting zones:

$$\lambda = \mathbf{211.65 + 55.09\boldsymbol{\alpha} + 0.1179\boldsymbol{\beta} + 0.0822\boldsymbol{\mu}} \tag{14}$$

For screw configuration with only conveying zones:

$$\lambda = \mathbf{89.78997 + 55.1197\boldsymbol{\alpha} + 0.11784\boldsymbol{\beta} + 0.08218\boldsymbol{\mu}} \tag{15}$$

where λ is the predicted mean residence time (s), α is the liquid to solid ratio, β is the screw speed (rpm), and μ is the powder flow rate (g/h). The linear ANN structure of this model is represented with 4 linear nodes, 2 in the first hidden layer and 2 in the second hidden layer; the first layer (process parameters) and the last layer (mean residence time) are related as in the above system of equations that accounts for all the nodes.

Figs. 7 & 8 show the contour profiles of the predicted mean residence time (s) with respect to two process parameters in each plot. In comparing plots, it is obvious that the kriging interpolation improved the smoothness of the contour profiles in all cases; while plots in Fig. 7 show so much rigidity in the change of the predicted mean residence time with the change of the process parameters (L/S ratio, screw speed, and powder flow rate) for all screw configurations, plots in Fig. 8 show more uniformity and smoothness with this change. The improvement in the smoothness of the predicted mean residence time with respect to the process parameters is mainly due to the decrease in the weight of the outliers and the increase in the weight of the grouped experimental points. Also, kriging interpolation widens the boundaries of the contour profiles, thus widening the prediction of the mean residence time at different interpolated process parameter points. This is due to the large number of the interpolated points that falls onto new points in the range of the process parameters.

Furthermore, in analysing the effect of the process parameters on the range of the mean residence time, it can be seen that screw configuration is the most influential parameter where the ranges of the mean residence time dramatically changed from 60-130 s for conveying elements to 100-160 s for 1 kneading zone, to 150-200 s for 2 kneading zones, to 190-250 s for 2 kneading zones and 1 cutting element. The conveying element geometric structure allow particles to be transported much more efficiently than kneading and cutting elements thus giving a shorter residence time in the granulator.

Fig. 7: Contours of mean residence times as function of process parameters and screw configuration before kriging.

Fig. 8: Contours of mean residence times as function of process parameters and screw configuration after kriging.

5. Conclusions

In this study, experimental data were collected and calculated for the mean residence time of granules produced in a twin-screw granulator at different process parameters and screw configurations. In order to develop an improved ANN model for this system, more data points are required, thus kriging interpolation was used to provide a greater range of data points from a limited number of measured data. The models were developed considering the effect of process parameters (*L/S* ratio, and screw speed) and screw configuration. The ANN model that was based on the kriging interpolated data showed a dramatic improvement in prediction accuracy (R^2, RMSE, and mean absolute deviation) in training and validating the ANN. This improvement was also reflected in the profiles where the derivatives got smooth and the prediction range got wide. The improved ANN model was derived in a set of linear equations that predict the mean residence time in terms of the process parameters and screw configuration. Better smoothness in the contour profiles was also obtained after kriging was implemented. From the contour profiles, it can be concluded that the screw configuration has the most important influence on the mean residence time. The improved ANN-kriging model proved to be very efficient in accurately predicting the mean residence time of twin-screw granulation in terms of the process parameters and screw configuration. The derived system of linear equations can further be used as a state space for a model predictive controller.

Further work is needed on the use of ANN and kriging interpolation for predicting more granule properties (particle size distribution, porosity, etc.) in terms of process parameters. Also, these empirical modelling techniques can be used to model other pharmaceutical manufacturing processes, such as mixing, tableting, and powder feeding. Further ANN-kriging models can be developed to take into account not just for process parameters but also material properties.

Appendix A

Interpolated mean residence times at new process parameters and screw configuration points are provided in a supplementary excel sheet. Also, screw configurations can be found in the supplementary file.

Abbreviations and Symbols

Symbols

a	coefficient used in Eq. 1
b	coefficient used in Eq. 1
c	coefficient used in Eq. 1
E	dye intensity
f	predicted value
F	activation function
h	Euclidian distance
L/S	liquid to solid ratio
K	number of experimental subsets used for ANN validation

n		number of experiments/points
R^2		coefficient of determination
t	[s]	time
$\bar{t}$	[s]	mean residence time
x		linear combination of input factors
x_k		interpolated point
y		measured value
y_k		response parameter
w		kriging weight

Subscript

i	experiment set

Acknowledgments

This work was funded by Science Foundation Ireland (SFI) under the following grants:

Model Predictive Control of Continuous Pharmaceutical Processes - (13/IA/1980).

The equipment used in this work was funded by Science Foundation Ireland (SFI) under the following grants:

Model Predictive Control of Continuous Pharmaceutical Processes - (13/IA/1980).

This research was made possible with the support of Pharmaceutical Manufacturing Technology Centre (PMTC) as well as Synthesis and Solid State Pharmaceutical Centre (SSPC).

This research was made possible by a grant from the Qatar National Research Fund under its National Priorities Research Program award number NPRP 10-0107-170119. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Qatar National Research Fund.

References

[1] C. Mangwandi, L. JiangTao, A.B. Albadarin, R.M. Dhenge, G.M. Walker, High shear granulation of binary mixtures: Effect of powder composition on granule properties, Powder Technology, 270 (2015) 424-434.
[2] S.L. Lee, T.F. O'Connor, X. Yang, C.N. Cruz, S. Chatterjee, R.D. Madurawe, C.M.V. Moore, L.X. Yu, J. Woodcock, Modernizing Pharmaceutical Manufacturing: from Batch to Continuous Production, Journal of Pharmaceutical Innovation, 10 (2015) 191-199.
[3] M. Sajjia, S. Shirazian, D. Egan, J. Iqbal, A.B. Albadarin, M. Southern, G. Walker, Mechanistic modelling of industrial-scale roller compactor 'Freund TF-MINI model', Computers & Chemical Engineering, 104 (2017) 141-150.
[4] M. Sajjia, S. Shirazian, C.B. Kelly, A.B. Albadarin, G. Walker, ANN Analysis of a Roller Compaction Process in the Pharmaceutical Industry, Chemical Engineering and Technology, (2017).
[5] S. Shirazian, M. Kuhs, S. Darwish, D. Croker, G.M. Walker, Artificial neural network modelling of continuous wet granulation using a twin-screw extruder, International Journal of Pharmaceutics.
[6] G.K. Reynolds, Modelling of pharmaceutical granule size reduction in a conical screen mill, Chemical Engineering Journal, 164 (2010) 383-392.
[7] T.C. Seem, N.A. Rowson, A. Ingram, Z. Huang, S. Yu, M. de Matas, I. Gabbott, G.K. Reynolds, Twin screw granulation — A literature review, Powder Technology, 276 (2015) 89-102.
[8] P. Suresh, I. Sreedhar, R. Vaidhiswaran, A. Venugopal, A comprehensive review on process and engineering aspects of pharmaceutical wet granulation, Chemical Engineering Journal, 328 (2017) 785-815.
[9] D. Ramkrishna, Population Balances Theory and Applications to Particulate Systems in Engineering, ACADEMIC PRESS, USA, 2000.
[10] A. Rogers, A. Hashemi, M. Ierapetritou, Modeling of Particulate Processes for the Continuous Manufacture of Solid-Based Pharmaceutical Dosage Forms, Processes, 1 (2013) 67.
[11] D. Barrasso, A. El Hagrasy, J.D. Litster, R. Ramachandran, Multi-dimensional population balance model development and validation for a twin screw granulation process, Powder Technology, 270, Part B (2015) 612-621.

[12] D. Barrasso, R. Ramachandran, Multi-scale modeling of granulation processes: Bi-directional coupling of PBM with DEM via collision frequencies, Chemical Engineering Research and Design, 93 (2015) 304-317.
[13] D. Barrasso, S. Walia, R. Ramachandran, Multi-component population balance modeling of continuous granulation processes: A parametric study and comparison with experimental trends, Powder Technology, 241 (2013) 85-97.
[14] A. Kumar, K.V. Gernaey, T.D. Beer, I. Nopens, Model-based analysis of high shear wet granulation from batch to continuous processes in pharmaceutical production – A critical review, European Journal of Pharmaceutics and Biopharmaceutics, 85 (2013) 814-832.
[15] A. Kumar, K.V. Gernaey, T. De Beer, I. Nopens, Model-based characterisation of twin-screw granulation system for continuous solid dosage manufacturing, in: J.K.H. Krist V. Gernaey, G. Rafiqul (Eds.) Computer Aided Chemical Engineering, Elsevier 2015, pp. 2165-2170.
[16] P. Kazemi, M.H. Khalid, J. Szlek, A. Mirtič, G.K. Reynolds, R. Jachowicz, A. Mendyk, Computational intelligence modeling of granule size distribution for oscillating milling, Powder Technology, 301 (2016) 1252-1258.
[17] M. Puri, A. Solanki, T. Padawer, S.M. Tipparaju, W.A. Moreno, Y. Pathak, Chapter 1 - Introduction to Artificial Neural Network (ANN) as a Predictive Tool for Drug Design, Discovery, Delivery, and Disposition: Basic Concepts and Modeling, Artificial Neural Network for Drug Design, Delivery and Disposition, Academic Press, Boston, 2016, pp. 3-13.
[18] F. Boukouvala, F.J. Muzzio, M.G. Ierapetritou, Dynamic Data-Driven Modeling of Pharmaceutical Processes, Industrial & Engineering Chemistry Research, 50 (2011) 6743-6754.
[19] F. Boukouvala, F.J. Muzzio, M.G. Ierapetritou, Design Space of Pharmaceutical Processes Using Data-Driven-Based Methods, Journal of Pharmaceutical Innovation, 5 (2010) 119-137.
[20] A. Kumar, M. Alakarjula, V. Vanhoorne, M. Toiviainen, F. De Leersnyder, J. Vercruysse, M. Juuti, J. Ketolainen, C. Vervaet, J.P. Remon, K.V. Gernaey, T. De Beer, I. Nopens, Linking granulation performance with residence time and granulation liquid distributions in twin-screw granulation: An experimental investigation, European Journal of Pharmaceutical Sciences, 90 (2016) 25-37.
[21] A. Kumar, M. Alakarjula, V. Vanhoorne, M. Toiviainen, F. De Leersnyder, J. Vercruysse, M. Juuti, J. Ketolainen, C. Vervaet, J.P. Remon, K.V. Gernaey, T. De Beer, I. Nopens, Linking granulation performance with residence time and granulation liquid distributions in twin-screw granulation: An experimental investigation, European Journal of Pharmaceutical Sciences, 90 (2016) 25-37.
[22] A. Kumar, G.M. Ganjyal, D.D. Jones, M.A. Hanna, Digital image processing for measurement of residence time distribution in a laboratory extruder, Journal of Food Engineering, 75 (2006) 237-244.
[23] B. Mu, M.R. Thompson, Examining the mechanics of granulation with a hot melt binder in a twin-screw extruder, Chemical Engineering Science, 81 (2012) 46-56.
[24] A. Poulesquen, B. Vergnes, P. Cassagnau, A. Michel, O.S. Carneiro, J.A. Covas, A study of residence time distribution in co-rotating twin-screw extruders. Part II: Experimental validation, Polymer Engineering & Science, 43 (2003) 1849-1862.

Table 1: Design of experiments for wet granulation using twin-screw granulator.

Run	*L/S*	Screw speed (rpm)	Powder flow rate (g/h)	Screw configuration
1	0.54	64	49.75	2 kneading zones
2	0.94	200	98.00	2 kneading zones
3	1.22	86	98.20	2 kneading zones
4	0.54	200	69.33	2 kneading zones
5	1.22	200	61.47	2 kneading zones
6	1.22	50	82.00	2 kneading zones
7	1.21	200	49.40	1 kneading zone
8	0.78	200	49.75	1 kneading zone
9	1.21	115	49.50	1 kneading zone
10	0.49	185	67.55	1 kneading zone
11	0.65	50	97.60	1 kneading zone
12	0.52	50	97.05	1 kneading zone
13	1.22	50	82.00	1 kneading zone
14	0.55	200	98.60	conveying elements only
15	1.22	200	98.15	conveying elements only
16	0.76	200	99.20	conveying elements only
17	0.59	85	51.00	conveying elements only
18	1.12	72	53.25	conveying elements only
19	0.48	50	49.80	conveying elements only
20	0.50	200	95.05	2 kneading zones with cutting elements
21	1.20	58	82.00	2 kneading zones with cutting elements
22	0.66	200	49.65	2 kneading zones with cutting elements
23	0.54	50	61.45	2 kneading zones with cutting elements
24	1.22	200	97.38	2 kneading zones with cutting elements

Table 2: Mean residence time values for various screw configurations.

2 kneading elements:							
run	1	2	3	4	5	6	
Mean residence time (s)	237.65	129.50	158.64	105.75	123.07	272.91	
One kneading elements:							
run	7	8	9	10	11	12	13
Mean residence time (s)	131.21	73.62	178.74	68.24	125.07	106.65	224.60
Conveying elements only:							
run	14	15	16	17	18	19	
Mean residence time (s)	18.41	127.52	44.82	87.79	148.62	91.41	
2 kneading zones with cutting elements:							
run	20	21	22	23	24		
Mean residence time (s)	99.74	277.24	222.89	250.95	222.84		

Table 3: Fitting parameters of training and validation of developed ANN.

Measures	Training original exp. data	Validating original exp. data	Training interpolated data	Validating interpolated data
R^2	0.8901	0.7581	0.9229	0.9242
RMSE	23.4378	36.1396	14.0317	13.9138
Mean Abs Dev	17.6228	26.4158	10.6985	10.6550
-LogLikehood	73.1726	40.0506	100273.2	50031.9
SSE	8789.32	10448.58	4862389.6	2390503
Sum Frequancy	16	8	24696	12348

Highlights

- Continuous wet granulation of pharmaceutical formulations
- ANN modelling of residence time distribution of granules in a twin-screw extruder
- Improving ANN predications by coupling with kriging model
- Study on effect of process parameters on residence time using the developed model

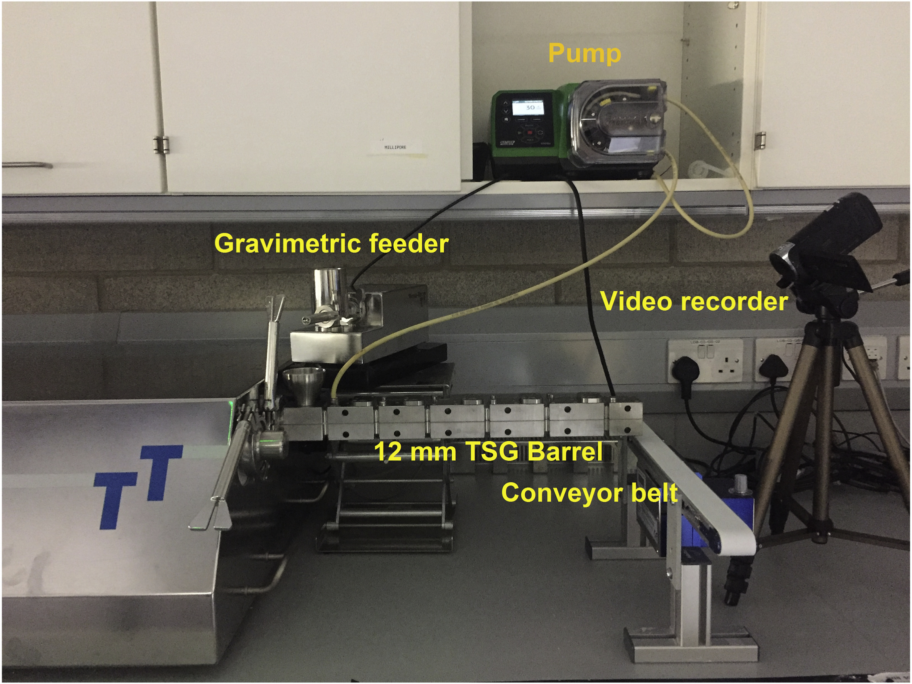
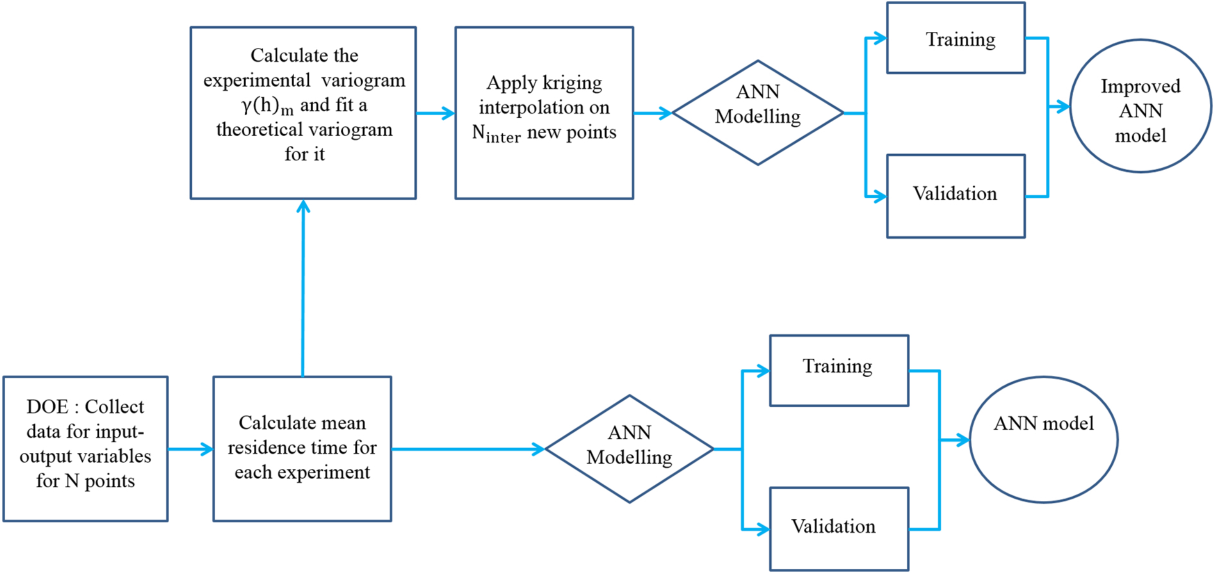

Pump
Gravimetric feeder
Video recorder
12 mm TSG Barrel
Conveyor belt

Figure 1



Calculate the experimental variogram γ(h)m and fit a theoretical variogram for it
Apply kriging interpolation on Ninter new points
ANN Modelling
Training
Validation
Improved ANN model
DOE : Collect data for input-output variables for N points
Calculate mean residence time for each experiment
ANN Modelling
Training
Validation
ANN model

Figure 2

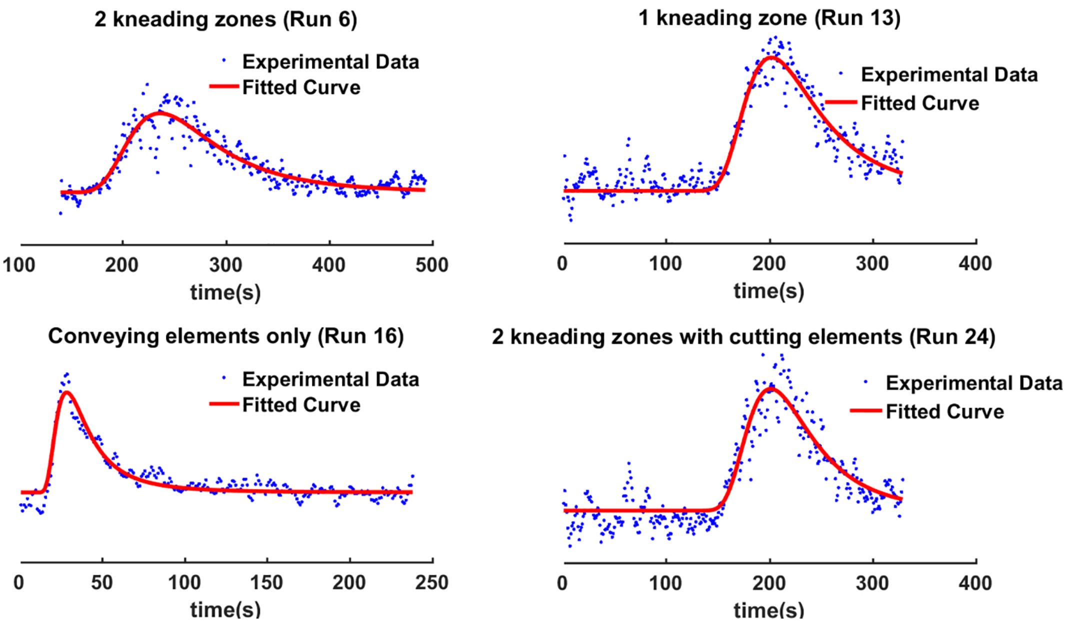

2 kneading zones (Run 6)
Experimental Data
Fitted Curve
100
200
300
400
500
time(s)
1 kneading zone (Run 13)
Experimental Data
Fitted Curve
0
100
200
300
400
time(s)
Conveying elements only (Run 16)
Experimental Data
Fitted Curve
0
50
100
150
200
250
time(s)
2 kneading zones with cutting elements (Run 24)
Experimental Data
Fitted Curve
0
100
200
300
400
time(s)

Figure 3

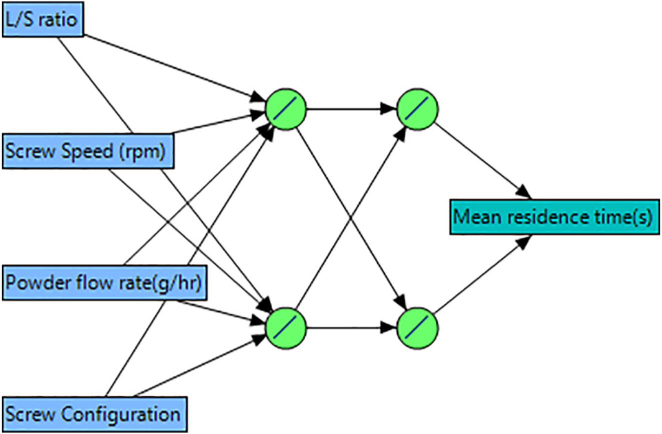
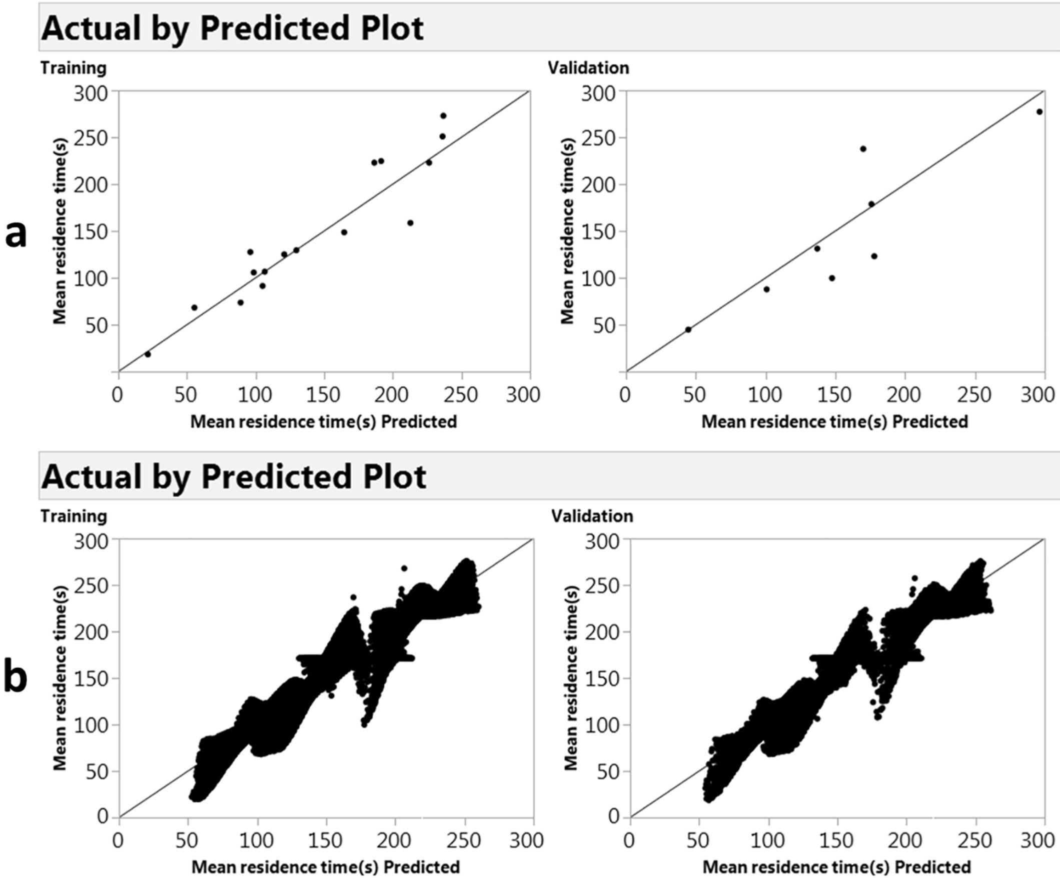
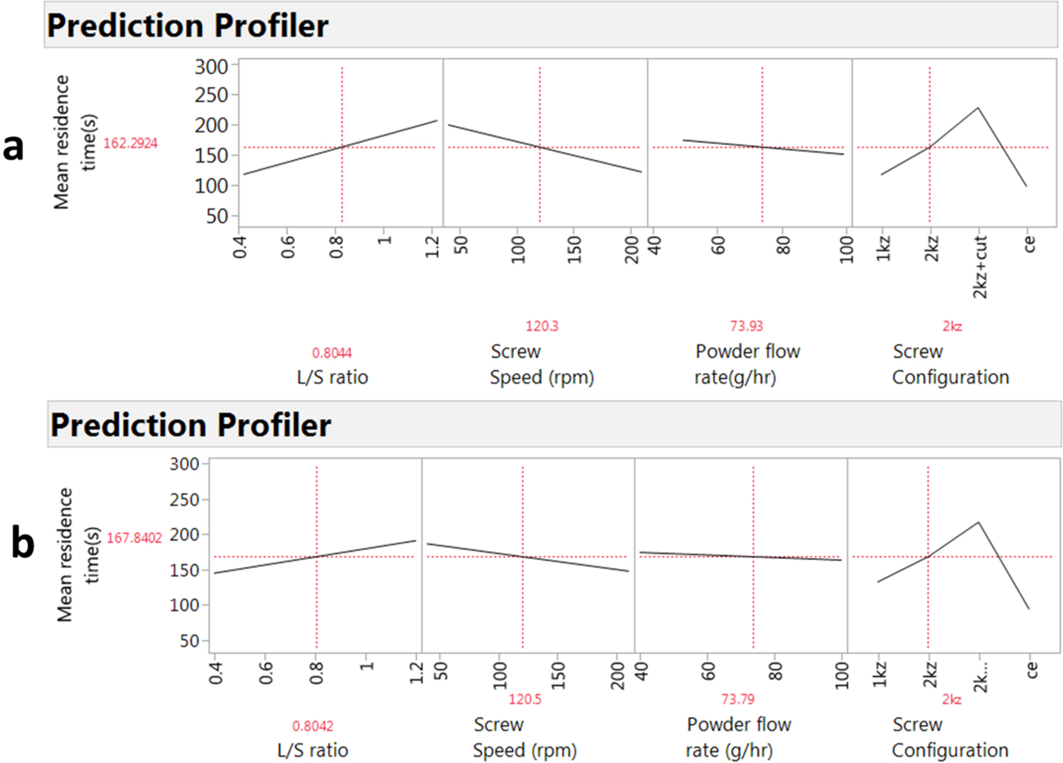

L/S ratio
Screw Speed (rpm)
Powder flow rate(g/hr)
Screw Configuration
Mean residence time(s)

Figure 4



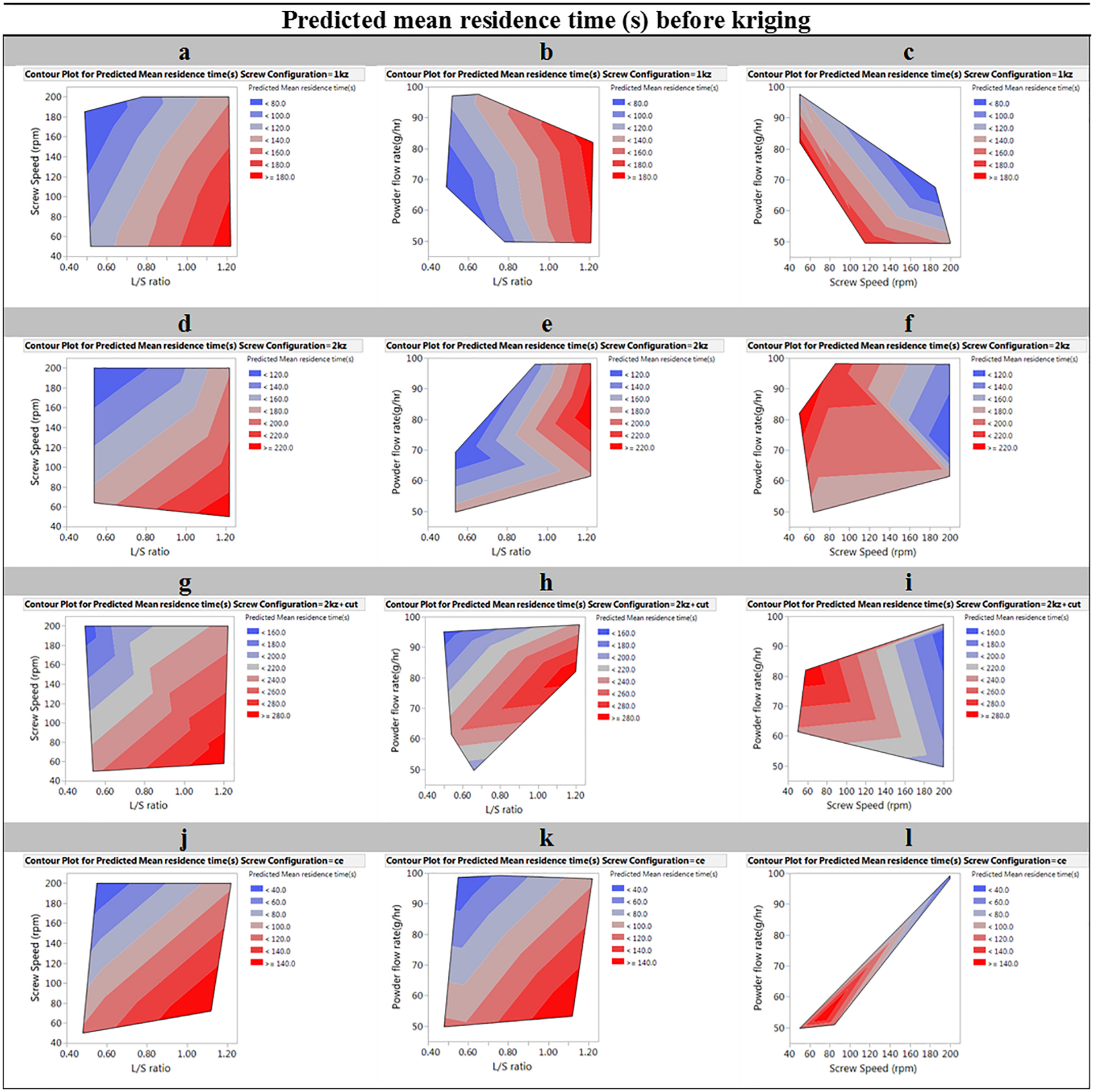
Actual by Predicted Plot
Training
Validation
a
Mean residence time(s)
Mean residence time(s) Predicted
300
250
200
150
100
50
0
Actual by Predicted Plot
Training
Validation
b

Figure 5



a
Prediction Profiler
Mean residence time(s)
162.2924
300
250
200
150
100
50
0.4
0.6
0.8
1
1.2
50
100
150
200
40
60
80
100
1kz
2kz
2kz+cut
ce
0.8044
L/S ratio
120.3
Screw Speed (rpm)
73.93
Powder flow rate(g/hr)
2kz
Screw Configuration
b
Prediction Profiler
Mean residence time(s)
167.8402
300
250
200
150
100
50
0.4
0.6
0.8
1
1.2
50
100
150
200
40
60
80
100
1kz
2kz
2k...
ce
0.8042
L/S ratio
120.5
Screw Speed (rpm)
73.79
Powder flow rate (g/hr)
2kz
Screw Configuration

Figure 6



Predicted mean residence time (s) before kriging
a
b
c
Contour Plot for Predicted Mean residence time(s) Screw Configuration=1kz
d
e
f
Contour Plot for Predicted Mean residence time(s) Screw Configuration=2kz
g
h
i
Contour Plot for Predicted Mean residence time(s) Screw Configuration=2kz+cut
j
k
l
Contour Plot for Predicted Mean residence time(s) Screw Configuration=ce
Predicted Mean residence time(s)
Screw Speed (rpm)
Powder flow rate(g/hr)
L/S ratio

Figure 7

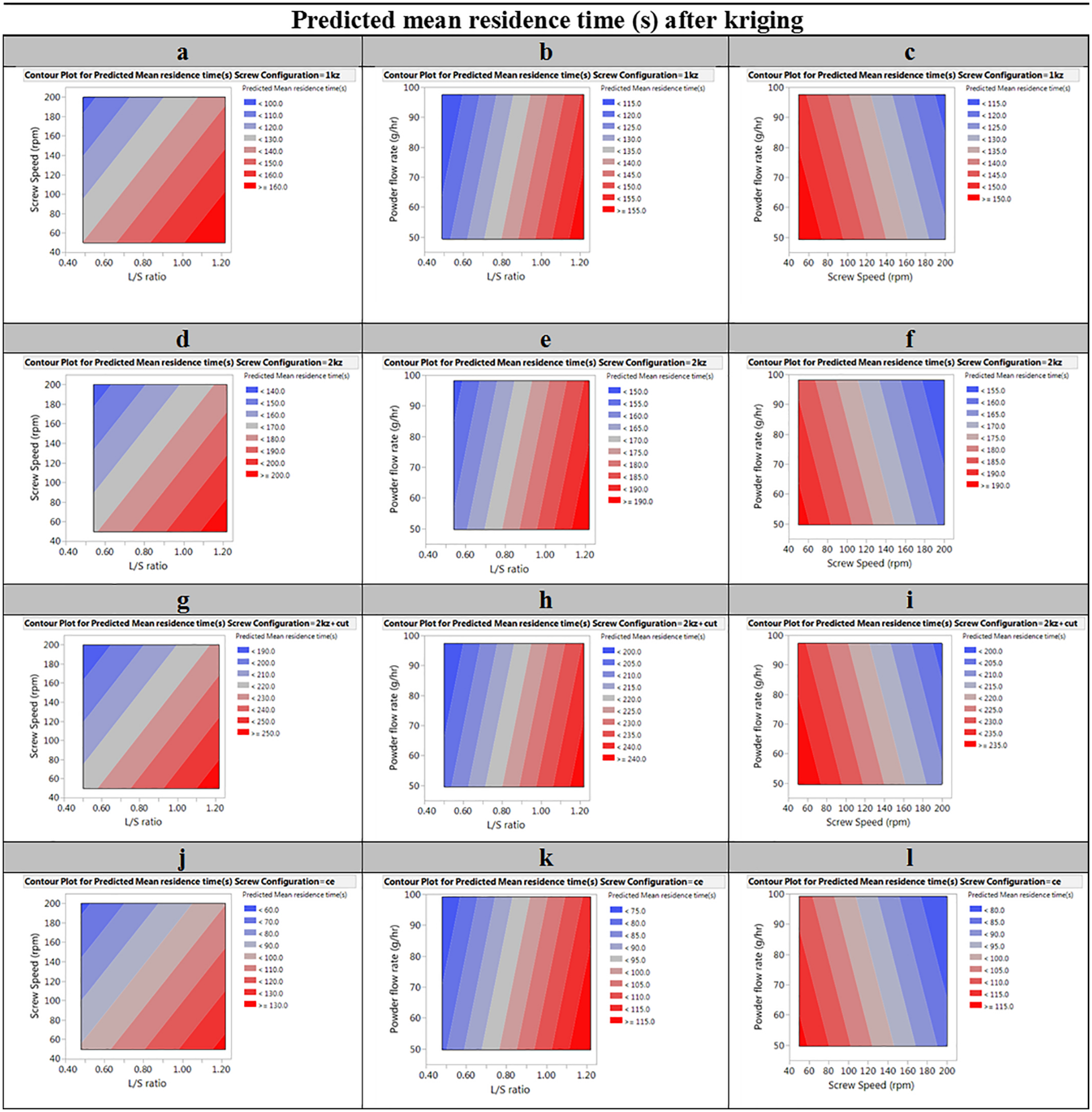

Predicted mean residence time (s) after kriging
a
Contour Plot for Predicted Mean residence time(s) Screw Configuration=1kz
Predicted Mean residence time(s)
< 100.0
< 110.0
< 120.0
< 130.0
< 140.0
< 150.0
< 160.0
>= 160.0
Screw Speed (rpm)
L/S ratio
b
Contour Plot for Predicted Mean residence time(s) Screw Configuration=1kz
Predicted Mean residence time(s)
< 115.0
< 120.0
< 125.0
< 130.0
< 135.0
< 140.0
< 145.0
< 150.0
< 155.0
>= 155.0
Powder flow rate (g/hr)
L/S ratio
c
Contour Plot for Predicted Mean residence time(s) Screw Configuration=1kz
Predicted Mean residence time(s)
< 115.0
< 120.0
< 125.0
< 130.0
< 135.0
< 140.0
< 145.0
< 150.0
>= 150.0
Powder flow rate (g/hr)
Screw Speed (rpm)
d
Contour Plot for Predicted Mean residence time(s) Screw Configuration=2kz
Predicted Mean residence time(s)
< 140.0
< 150.0
< 160.0
< 170.0
< 180.0
< 190.0
< 200.0
>= 200.0
Screw Speed (rpm)
L/S ratio
e
Contour Plot for Predicted Mean residence time(s) Screw Configuration=2kz
Predicted Mean residence time(s)
< 150.0
< 155.0
< 160.0
< 165.0
< 170.0
< 175.0
< 180.0
< 185.0
< 190.0
>= 190.0
Powder flow rate (g/hr)
L/S ratio
f
Contour Plot for Predicted Mean residence time(s) Screw Configuration=2kz
Predicted Mean residence time(s)
< 155.0
< 160.0
< 165.0
< 170.0
< 175.0
< 180.0
< 185.0
< 190.0
>= 190.0
Powder flow rate (g/hr)
Screw Speed (rpm)
g
Contour Plot for Predicted Mean residence time(s) Screw Configuration=2kz+cut
Predicted Mean residence time(s)
< 190.0
< 200.0
< 210.0
< 220.0
< 230.0
< 240.0
< 250.0
>= 250.0
Screw Speed (rpm)
L/S ratio
h
Contour Plot for Predicted Mean residence time(s) Screw Configuration=2kz+cut
Predicted Mean residence time(s)
< 200.0
< 205.0
< 210.0
< 215.0
< 220.0
< 225.0
< 230.0
< 235.0
< 240.0
>= 240.0
Powder flow rate (g/hr)
L/S ratio
i
Contour Plot for Predicted Mean residence time(s) Screw Configuration=2kz+cut
Predicted Mean residence time(s)
< 200.0
< 205.0
< 210.0
< 215.0
< 220.0
< 225.0
< 230.0
< 235.0
>= 235.0
Powder flow rate (g/hr)
Screw Speed (rpm)
j
Contour Plot for Predicted Mean residence time(s) Screw Configuration=ce
Predicted Mean residence time(s)
< 60.0
< 70.0
< 80.0
< 90.0
< 100.0
< 110.0
< 120.0
< 130.0
>= 130.0
Screw Speed (rpm)
L/S ratio
k
Contour Plot for Predicted Mean residence time(s) Screw Configuration=ce
Predicted Mean residence time(s)
< 75.0
< 80.0
< 85.0
< 90.0
< 95.0
< 100.0
< 105.0
< 110.0
< 115.0
>= 115.0
Powder flow rate (g/hr)
L/S ratio
l
Contour Plot for Predicted Mean residence time(s) Screw Configuration=ce
Predicted Mean residence time(s)
< 80.0
< 85.0
< 90.0
< 95.0
< 100.0
< 105.0
< 110.0
< 115.0
>= 115.0
Powder flow rate (g/hr)
Screw Speed (rpm)

Figure 8